AMENDMENTS TO SPECIFICATION

Please replace paragraph [0002] with the following amended paragraph:

[0001] Prior to the advent of VoiceXML (Voice Extensible Markup Language) and its precursor languages, VoxML, SpeechML, and ~~others. Speech~~ other, speech applications were described (or programmed) using standard programming techniques, e.g. C/C++ programs that made function (or object invocation) calls to lower level device drivers and speech recognition engines. For example, companies such as Nuance Communications, Inc., Menlo Park, California, and SpeechWorks International, Inc., Boston, Massachusetts, have developed sophisticated automated speech recognition (ASR) systems and provide complex C/C++ interfaces called software development kits (SDKs) to allow customers to develop systems.

Please replace paragraph [0008] with the following amended paragraph:

[0002] Thus, while subtraction of starttime and endtime JavaScript variables would result in a fairly good approximation of the time from the start of *all* audio playback for a given VoiceXML state and the entry into the next VoiceXML state, it will not be relative to the apparent position of the <var/> declaration in the code or the second prompt. Thus to perform any calculations about barge-in it would be necessary to know the playback time of all audio prompts for the previous VoiceXML state. This may be impossible to determine

in the interpreter if speed-adjusting technologies are used to increase playback speeds and reduce pauses between ~~words thus~~ words. Thus the apparent file size/sampling rate may not be the same as playback time.

Please replace paragraph [0053] with the following amended paragraph:

[0003] The problem of false barge in was already discussed briefly above. However, a fuller discussion is useful to consider. Although humans can do a relatively good job at comprehending other humans even in loud/noisy environments, speech recognition systems do not fare as well and when you add in the (poor) quality of many (wireless) telephone networks, the situation gets worse. Other factors such as road noise, stadium noise, bar noise, etc., ~~all makes~~ all make the problem worse. All of those noises might be considered by the speech recognition system as a cue that speech has started—a false barge in.

Please replace paragraph [0062] with the following amended paragraph:

[0004] Similarly, if a particular grammar has a large number of phonotactically similar options the strategy selected can be adjusted further by the application programmer. For example, a grammar of United States equity issues (stocks/company names/ticker symbols) is fairly large (thousands of options) with many phonotactically similar options. In such a case the starting strategy upon inferring that an error occurred might be the fourth approach, e.g.

"Sorry, when you hear the stock company name you want, say 'tell me more'... Cisco Corporation... Sysco Foods...".

Please replace paragraph [0064] with the following amended paragraph:

[0005] More specifically, at the point where the selection 310 is made a first grammar, "MenCollBasketballTeamChoices", would be active and then at a later point, e.g. when the cancel 320A (or 320B) came, a second grammar, "ScoreNavigationCommands", would be active. Since the second grammar does not include the options from the first grammar [[than]] then one of two things will happen if the user repeats a sports team name the speech recognizer will either: (i) false accept the team name as one of the options in the second grammar or (ii) correctly reject the team name as out of grammar, resulting in a <nomatch/>.